Short communication

NMR-based metabonomics of bovine blood:

An investigation into the effects of long term storage on plasma samples

Manuela Trabi [1#*], Marianne D. Keller [2,3#], Nicholas N. Jonsson[3,4]

[1,2,3]The University of Queensland, St. Lucia, Queensland 4072, Australia ([1]Centre for Integrative Complementary and Molecular Medicine, [2]Center for Advanced Imaging, [3]School of Veterinary Science); [4]The University of Glasgow, College of Medical, Veterinary and Life Sciences.

[#]These authors contributed equally.

Short title: Metabolite changes in long term storage of blood plasma

*To whom correspondence should be addressed: m.trabi@uq.edu.au

Current address for reprint requests and other correspondence:

Manuela Trabi

University of Queensland Diamantina Institute

Level 4 Research Wing, Princess Alexandra Hospital

Woolloongabba, Queensland, 4102

Australia

Abstract

Freezers in research institutions often contain a plethora of samples left over from studies performed years or even decades ago. Along with samples stored in biobanks, these could prove to be treasure troves for metabonomic research. Although the influence of sample handling and short to medium term storage on conventionally determined blood parameters has been reported, little is known about the effects of long term storage (years to decades) on plasma samples. The aim of this study was to investigate the influence of long term storage on the metabolite profile and to assess the value of archived samples for metabonomic studies. Heparinised plasma samples of 22 Holstein Friesian heifers that had been stored at −20° C for between 2 and 15 years were analysed using proton NMR spectroscopy and statistical analysis techniques. Lactate accounted for principal component one and explained 79.6% of variance between all spectra, but was not correlated with storage time. The highest correlation with storage time (R^2 = 0.474) was found for betaine, with other metabolites (acetoacetate, histidines, glycerol, lipids and glucose) also showing moderate correlation (R^2 values between 0.217 and 0.437). Our results indicate that samples stored for extended periods of time can potentially be used in metabolomics studies, if precautions are taken during data analysis.

Keywords: betaine, lactate, principal component analysis, metabolomics, cattle, metabolite stability.

Abbreviations: NMR: nuclear magnetic resonance; PCA: principal component analysis; PC: principal component; standard three letter codes were used for amino acids.

1 Introduction

Metabonomic analysis of plasma samples, both by NMR and mass spectrometry based methods, has become increasingly common over the last decade. As with every analytical endeavor, sample quality is crucial for success. Handling and storage of blood samples can significantly change their metabolite profiles and potentially generate misleading results, yet comparatively few studies have been published on the effects of handling and in particular storage of blood samples (*i.e.* serum and plasma) used in a metabonomics context. Several authors have described changes in the concentration of up to 15 metabolites when serum was left in contact with the blood clot for extended periods of time (see *e.g.* Teahan *et al.* 2006; Boyanton and Blick 2002).

In contrast, the short to medium term storage of plasma or serum, once it has been separated from the cellular matter, seems to only have a minor influence on the metabolite profile. Negligible changes in the ^{1}H NMR spectra of plasma samples stored for up to 6 days at +4° C (Otvos *et al.* 1991) or up to 9 months at −80° C (Deprez *et al.* 2002) have been reported, but a decrease in triglyceride resonances after storage or with an increased number of freeze-thaw cycles has also been described (Bell *et al.* 1987; Bell *et al.* 1988). Brinc *et al.* (2012) detected no significant time-dependent concentration changes for the majority of biochemical markers tested over up to 13 months, while other authors reported changes in amino acid concentrations after up to 6 months (Schaefer *et al.* 1987; van Eijk *et al.* 1994).

To our knowledge, prior metabonomic studies investigating the effects of sample storage did not extend beyond 9 months. The current study compares the metabolite profiles of bovine plasma

stored at −20° C for between 2 and 15 years using ^{1}H NMR techniques and multivariate data analysis. The metabolite profiles were largely stable, indicating that the pool of samples available for metabonomics studies can potentially be expanded to those stored for an extended period of time.

2 Materials and methods

2.1 Sample collectionSamples had been obtained from 22 clinically normal, 6–12 month old Holstein Friesian heifers for the purpose of parentage testing. Blood had been collected into heparinised tubes, centrifuged and the plasma had been stored at −20° C for between 2 and 15 years (y) (2×2y, 1×3y, 1×7y, 1×8y, 13×10y, 1×12y, 1×13y, 1×14y, 1×15y).

2.2 NMR spectroscopy

All plasma samples were thawed at room temperature immediately before analysis. If the available sample volume was below 500 μL, bidest H_2O was added and the resulting dilution effect was compensated for during data analysis by normalization of signal intensities. Finally, 50 μL $D_2O$for NMR frequency lock was added. One-dimensional CPMG spectra with 0.5ms T_2 spin echo time, a 3s relaxation delay between successive scans (256 scans per spectrum, 16K data points, spectral width 6009.6 Hz) and selective water presaturation during the relaxation delay were recorded using an Avance 500MHz spectrometer (Bruker BioSpin, Karlsruhe, Germany). Prior to Fourier transformation, an exponential multiplication window function (line-broadening factor 1Hz) was applied. Phase and baseline correction of all spectra was performed manually. Sodium 2,2-dimethylsilapentane-5-sulphonate (main signal at 0ppm) was added to one sample for chemical shift referencing and all spectra were subsequently referenced using the relative chemical shift of the CH_3

group of lactate (1.33ppm).

2.3 Data reduction and analysis

Data analysis was performed in AMIX 3.6.8 (Bruker BioSpin, Germany). The chemical shift region from 0.7 to 8.6ppm of each spectrum was devided into buckets (spectral integral regions) of 0.04ppm, initially excluding only the residual water signal (4.40 to 5.20ppm), but in subsequent rounds of analysis also spectral regions containing lactate signals (1.29-1.38ppm and 4.088-4.147ppm). Integral intensities for each spectrum were normalized to the total sum of integrals and the resulting bucket tables were subjected to basic principal component analysis (PCA) without further scaling.

Metabolites were identified using databases (BBIOREFCODE, Bruker BioSpin, Germany; Human Metabolome Database, http://www.hmdb.ca; Biological Magnetic Resonance Databank, http://www.bmrb.wisc.edu; Magnetic Resonance Metabolomics Database, http://mdl.imv.liu.se) as well as published data. For statistical analysis of selected metabolites, spectral regions (listed in Supplementary Table S-1) were chosen manually, integrated and again normalized to allow for direct comparison of the samples.

3 Results and discussion

3.1 Analysisof ^{1}H NMR profiles

Upon visual inspection, the spectra did not show any obvious features that would indicate a relation to storage time (see Supplementary Figure S-1 for representative spectra). We employed PCA, an unsupervised multivariate statistical technique, to detect subtle patterns potentially present in our

dataset. When only the residual water signal was excluded from analysis, the influence plot showed that all spectra were well inside the 95% confidence limits, and the bucket containing the CH_3 doublet of lactate (1.34ppm) was identified as PC1 (Fig. 1). However, our data did not indicate a significant correlation of PC1 with storage time (R^2 = 0.0203; Supplementary Figure S-2;). Lactate is one of the most variable metabolites in both human and rat plasma and in ruminants is largely attributable to the ruminal microflora and diet. This natural fluctuation as well as handling of blood samples prior to storage probably explain the variation in lactate peak integral values seen here.

When spectral regions containing lactate signals were excluded from analysis in order to evaluate the influence of other metabolites on PCA, PC1 accounted for 32.96% of variance (Fig. 2a), but again no strict clustering of samples according to storage time was observed. Buckets contributing most to PC1 were 3.42ppm, 3.46ppm (glucose peaks, but also peaks of taurine, Pro and Thr) and 1.46ppm (mainly Ala) (Fig. 2b). No significant correlation of any of these buckets with storage time could be established (Supplementary Figure S-3; R^2 values between 0.0859 and 0.1580).

3.2 Correlation analysis of buckets

A Kolmogorov-Smirnov test of our data matrix confirmed that all buckets followed a normal distribution, making a simple correlation analysis feasible. Tab. 1 lists all buckets with $|r| > 0.423$, the critical value of the Pearson Product-Moment correlation coefficient for a dataset of n = 22 and p = 0.05 (Rohlf and Sokal 1995).

The highest correlation was found for bucket 3.26ppm (r = -0.625, R^2 = 0.391), which is dominated by a betaine (trimethylglycine) peak. Although the p value for this bucket is very low (0.002), there is a moderate negative correlation with storage time, as reflected by the r and R^2 values (Tab. 1). Intensities

for this bucket vary considerably in samples stored for between 2 and 4 years (Supplementary Figure S-4), possibly reflecting differences in betaine intake of cattle due to different feed. Four buckets (2.30, 1.58, 1.54 and 1.42ppm) showed a moderate negative correlation to bucket 3.26ppm, with r values between -0.467 and -0.574. None of these buckets contained signals of dimethylglycine or glycine, the *in vivo* metabolites of betaine. If the decreased betaine levels seen in the older samples were due solely to a simple degradation of betaine, a higher negative correlation of bucket 3.26ppm with another bucket would be expected – unless of course betaine is converted to an analyte that cannot be detected with the NMR techniques employed here.

Table 1: Buckets with correlations to storage time significantly different from random (critical cutoff value 0.423) and their corresponding significance values and coefficients of determination when only the residual water signal was excluded. Compounds in brackets contribute minor peaks to the bucket integral. Scatter plots for all buckets with $|r| > 0.5$ (*i.e.* $R^2 > 0.25$) are given in Supplementary Figure S-4.

bucket	r	p	R^2	putative assignments
3.26	-0.625	0.002	0.391	betaine, (glucose, trimethylamine-N-oxide, Phe)
2.22	-0.594	0.004	0.353	acetoacetate, (3-hydroxy butyrate)
3.58	0.550	<0.050	0.302	glycerol, (Val)
0.82	-0.547	0.008	0.300	cholesterol, lipids
4.18	0.505	0.017	0.255	3-hydroxybutyrate
6.86	-0.504	0.017	0.254	Tyr
1.42	0.485	0.022	0.235	Ala
7.70	-0.475	0.026	0.225	imidazole ring
2.42	0.470	0.027	0.221	Gln
7.78	0.449	0.036	0.201	imidazole ring
7.42	0.446	0.037	0.199	Phe
5.38	-0.429	0.047	0.184	lipid
7.02	-0.427	0.047	0.182	imidazole ring
3.78	0.424	0.049	0.180	glucose, glycerol, Ala

r: correlation; p: significance; R^2: coefficient of determination

3.3 Statistical analysis of peak integrals

Rectangular bucketing as used in our initial analyses can result in buckets containing peaks of several metabolites or a metabolite peak being represented in two neighbouring buckets. We therefore superimposed all spectra and manually chose spectral regions (listed in Supplementary Table S-1) that covered the target resonances of selected metabolites, but excluded interfering peaks.

After integration and normalization, all values were again normally distributed. Eight integrals with $|r| > 0.423$ (the critical value of the Pearson Product-Moment correlation coefficient, as outlined above) were identified (Supplementary Table S-2). Most of these integrals showed a negative correlation with storage time, with the exception of glycerol and glucose, which were positively correlated. The integral with the highest correlation was Integral 25, encompassing a spectral area from 3.25-3.28ppm, *i.e.* most of the region covered by bucket 3.26ppm identified above (Tab. 1). Integral 25, however, only contained the peak arising from the CH_3 groups of betaine. With $r = -0.688$ ($R^2 = 0.474$), the correlation identified here is again a moderate one. The other integral (Int.17) containing a betaine peak (arising from the CH_2 group) showed a smaller correlation ($r = -0.401$). Betaine is an important cofactor in the methylation cycle and in humans and animals is a product of choline oxidation or originates from nutritional sources. Changes in cattle feed amounts and composition over the last decades as well as differences in pasture composition due to pasture improvemen might have influenced the differences in betaine concentration seen here, rather that the extended storage time. Indeed, Hustad *et al.* (2012) found betaine concentrations to be stable for up to 29 years in human serum.

The methyl signal of acetoacetate (Int. 37) showed the second highest correlation to storage time ($r = -0.661$). Rapid non-enzymatic conversion of acetoacetate, has been reported previously (Carragher *et al.* 2003). In serum samples, estimates for losses upon storage range from less than 2% (Työppönnen and Kauppinen 1980) to 40% in one week (Fritzsche *et al.* 2001). Fritzsche *et al.* (2001) suggested that in serum samples stored at −20° C, virtually all acetoacetate was degraded after 40 days, which was clearly not the case for our samples, possibly influenced by the fact that we used heparinised plasma, not serum.

Integrals 3 and 6 (imidazole compounds, *i.e.* His, 1-methyl-His, 3-methylHis) showed weak to moderate negative correlation with storage time. Hubbard and Mejia (1995) reported that the concentrations of these compounds remained stable in serum samples stored at −70° C for up to 4 weeks.

Integral 23 (glycerol) showed a moderate positive correlation with storage time. Glycerol is a degradation product of triglycerides, the major components of very low density lipoproteins (VLDL), and blood glycerol levels are therefore linked to lipolysis *in vivo* and *in vitro*. In accordance with this, Integral 46, containing peaks attributed mainly to lipids from low density lipoproteins and VLDL, but also cholesterol (Nicholson *et al.* 1995), showed a moderate negative correlation with storage time. The effect of sample storage on lipid fractions is a contentious one, with some authors reporting essentially no effect on triglyceride levels for up to 10 years (Matthan *et al.* 2010), but others finding increases in VLDL triglyceride and phospholipid concentrations after just 10 days

(Evans *et al.* 1995). The concentration of free glycerol itself has been reported to be stable for at least 96h in serum stored at 4° C (Stinshoff *et al.* 1977) – With relation to the samples used here, it is conceivable that the average free plasma glycerol concentration has decreased over the years due to changes in cattle management, since free glycerol is mainly generated when the body uses stored fat as a source of energy. Improved year round nutrition would discourage the utilization of fat deposits, leading to lower free glycerol concentrations in the blood.

Interestingly, Integral 20 (glucose) showed a weak positive correlation to storage time, whereas in the literature, a decrease of glucose concentration in the order of 10% in 24h to 14 months (Clark *et al.* 1990) or 15% in 19 months (Giampietro *et al.* 1980) has been reported.

3.4 Summary

In summary, we have shown here that in this set of samples, weak to moderate correlations exist between storage time and the levels of a number of metabolites. However, it stands to reason that for most metabolites, natural variation as well as variation introduced through sample handling overshadow the effects of long term storage.

A variety of external factors, including sampling techniques and handling prior to storage, can affect the analyte levels in serum and plasma samples. It can for example be assumed that the older samples in this study had been collected into glass containers, a material with adsorption characteristics considerably different to those of the modern plastic containers. . It is also highly likely that tubes from different manufacturers were used, possibly with different concentrations of heparin, an anticoagulant that has been reported to decrease glucose, creatinine and cholesterol

concentrations (Gelfert and Staufenbiel 1998).

Plasma samples used in this study were originally collected for parentage testing and no details of sample handling before freezing are known . In veterinary practice, it is quite common for blood samples not to be cooled or centrifuged immediately after collection. However, since this practice is unlikely to have changed significantly in the last 15 years, a correlation with the age of the sample is unlikely.

Long term storage harbours an increased risk of freezing and thawing cycles due to system malfunction. Changes in the concentration of NMR observable metabolites through freezing and thawing are controversial, with both a statistically significant decrease in glucose, cholesterol and volatile fatty acids and an increase in lactate reported by one research team (Laborde *et al.* 1995), but no influence found by another (Morris *et al.* 2002). Repetitive freezing and thawing of plasma has also been reported to decrease the concentration of urea-N (Deane *et al.* 2004). In our dataset, however, urea did not show any significant correlation to storage time ($r = -0.06$). A further side effect of long term storage is potential sublimation of the matrix and therefore concentration of the analytes. Normalization of integral intensities to the total sum of integrals for each spectrum, as employed in this study, largely eliminates the confounding influence of sublimation and other concentration effects that affect the whole sample, like sample dilution or hydration status of the test subject.

Samples stored for less than two years were not included in this study, since multiple publications on short and medium-term storage are available (*e.g.* Brinc *et al.* 2012; Deprez *et al.* 2002; van Eijk *et al.* 1994) Research into the effects of long term storage of plasma or serum samples is sparse (*e.g.* Hustad *et al.* 2012; Gislefoss *et al.* 2008) or, in the context of metabolomics, non-existent.

Investigations into the effects of time on chemical compounds and mixtures pose a challenge with today's short research funding cycles. Analysis of aliquots of the same samples stored for several years would have been preferable – albeit unrealistic. A study into the storage of blood samples obtained from animals of the same breed, sex and age (prior to reproduction), as performed here, is in all probability the next best option.

4 Concluding remarks

This pilot study indicates that long term storage of heparinised plasma samples for up to 15 years potentially only affects a small number of metabolites. However, due to the small number of samples in this study, the results presented here should be regarded as preliminary. Although bovine plasma was used in this study, similar results would most likely be obtained with human plasma, given the high similarity of metabolite profiles of human and bovine blood. The possibility of using samples that have been stored for extended periods of time in metabonomics studies will potentially give researchers access to a larger pool of material, for example for investigations into rare diseases. Archived samples also provide a window back in time, opening up for example the intriguing possibility of investigations into the effect of lifestyle changes on modern diseases.

Acknowledgements The authors thank Dr. U.C. Marx (IMB, UQ) and A. Daley (AGL, UQ) for their help and contributions.

Electronic supplementary material: The online version of this article (doi: xxx) contains supplementary material which is available to authorized users.

References:

Bell, J. D., Brown, J. C. C., Norman, R. E., Sadler, P. J. & Newell, D. R. (1988). Factors affecting ^{1}H NMR spectra of blood plasma: cancer, diet and freezing. *NMR in Biomedicine*, 1, 90-94.

Bell, J. D., Sadler, P. J., Macleod, A. F., Turner, P. R. & La Ville, A. (1987). ^{1}H-NMR studies of human blood plasma - assignment of resonances for lipoporteins. *FEBS Letters*, 219, 239-243.

Boyanton, B. L. J. & Blick, K. E. (2002). Stability studies of twenty-four analytes in human plasma and serum. *Clinical Chemistry*, 48, 2242-2247.

Brinc, D., Chan, M. K., Venner, A. A., *et al*. (2012). Long-term stability of biochemical markers in pediatric serum specimens stored at -80 degrees C: A CALIPER Substudy. *Clinical Biochemistry*, 45, 816-826.

Carragher, F. M., Bonham, J. R. & Smith, J. M. (2003). Pitfalls in the measurement of some intermediary metabolites. *Annals of Clinical Biochemistry*, 40, 313-320.

Clark, M. L., Humphreys, S. M. & Frayn, K. N. (1990). Stability of plasma glucose during storage. *Annals of Clinical Biochemistry*, 27, 373-377.

Deane, J. C., Dagleish, M. P., Benamou, A. E. M., Wolf, B. T. & Marlin, D. (2004). Effects of syringe material and temperature and duration of storage on the stability of equine arterial blood gas variables. *Veterinary Anaesthesia and Analgesia*, 31, 250-257.

Deprez, S., Sweatman, B. C., Connor, S. C., Haselden, J. N. & Waterfield, C. J. (2002). Optimisation of collection, storage and preparation of rat plasma for ^{1}H NMR spectroscopic analysis in toxicology studies to determine inherent variation in biochemical profiles. *Journal of Pharmaceutical and Biomedical Analysis*, 30, 1297-1310.

Evans, K., Mitcheson, J. & Laker, M. F. (1995). Effect of storage at 4° C and -20° C on lipid, lipoprotein and apolipoprotein concentrations. *Clinical Chemistry,* 41, 392-396.

Fritzsche, I., Bührdel, P, Melcher, R. & Böhme, H. J. (2001). Stability of ketone bodies in serum in dependence on storage time and storage temperature. *Clinical Laboratory*, 47, 399-403.

Gelfert, C. C. & Staufenbiel, R. (1998). Proper sampling and handling of blood samples to secure reliable results for the diagnosis of metabolic disturbances in cattle. *Der Praktische Tierarzt*, 79, 640-650.

Giampietro, O., Navalesi, R., Buzzigoli, G., Boni, C. & Benzi, L. (1980). Decrease in plasma glucose concentration during storage at -20° C. *Clinical Chemistry,* 26, 1710-1712.

Gislefoss, R. E., Grimsrud, T. K. & Mørkrid, L. (2008). Long-term stability of serum components in the Janus Serum Bank. *The Scandinavian Journal of Clinical & Laboratory Investigation*, 68, 402-409.

Hubbard, R. W. & Mejia, A. (1995). Human plasma preservation for amino acid analysis by immediate low-temperature freezing. *Clinical Biochemistry*, 28, 318-320.

Hustad, S., Eussen, S., Øivind, M., *et al.* (2012). Kinetic modeling of storage effects on biomarkers related to B vitamin status and one-carbon metabolism. *Clinical Chemistry*, 58, 402-410.

Laborde, C. J., Chapa, A. M., Burleigh, D. W., Salgado, D. J. & Fernandez, J. M. (1995). Effects of processing and storage on the measurement of nitrogenous compounds in ovine blood. *Small Ruminant Research*, 17, 159-166.

Matthan, N. R., Ip, B., Resteghini, N., Ausman, L. M. & Lichtenstein, A. H. (2010). Long-term fatty acid stability in human serum cholesteryl ester, triglyceride, and phospholipid fractions.

Journal of Lipid Research, 51, 2826-2832.

Morris, J. D., Fernandez, J. M., Chapa, A. M., Gentry, L. R., Thorn, K. E. & Weick, T. M. (2002). Effects of sample handling, processing, storage, and hemolysis on measurements of key energy metabolites in ovine blood. *Small Ruminant Research*, 43, 157-166.

Nicholson, J. K., Foxall, P. J. D., Spraul, M., Farrant, R. D. & Lindon, J. C. (1995). 750MHZ ^{1}H and ^{1}H-^{13}C NMR spectroscopy of human blood plasma. *Journal of Analytical Chemistry*, 67, 793-811.

Otvos, J. D., Jeyarajah, E. J. & Bennett, D. W. (1991). Quantification of plasma lipoproteins by proton nuclear magnetic resonance spectroscopy. *Clinical Chemistry*, 37, 377-386.

Rohlf, F. J. & Sokal, R. R. (1995). *Statistical tables*, W. H. Freeman, New York.

Stinshoff, K., Weisshaar, D., Staehler, F., Hesse, D., Gruber, W. & Steier, E. (1977). Relation between concentrations of free glycerol and triglycerides in human sera. *Clinical Chemistry*, 23, 1029-1032.

Teahan, O., Gamble, S., Holmes, E., *et al.* (2006). Impact of analytical bias in metabonomic studies of human blood serum and plasma. *Journal of Analytical Chemistry*, 78, 4307-4318.

Työppönnen, J. & Kauppinen, K. (1980). The stabiliy and automatic determination of ketone bodies in blood samples taken in field conditions. *Acta Veterinaria Scandinavica*, 21, 55-61.

van Eijk, H. M. H., Dejong, C. H., Deutz N. E. & Soeters, P. B. (1994). Influence of storage conditions on normal plasma amino-acid concentrations. *Clinical Nutrition*, 13, 374-380.

Figure legends

Fig. 1: Scores plot (a) and loadings plot (b) of PC1 *vs.* PC2 when only the residual water signal was excluded. The storage time is indicated in the scores plot by the symbol (◆: less than 10y; ◆: 10-11y; ◇: more than 11y) and given next to each symbol (in years. months as a fraction of years). Numbers in the loadings plot correspond to the chemical shift (in ppm) at the centre of the bucket.

Fig. 2: Scores plot (a) and loadings plot (b) of PC1 *vs.* PC2 when both the residual water signal and spectral regions containing lactate signals were excluded. For explanation of symbols see Fig. 1.

Supplementary material:

Figure legends:

Fig. S-1: ^{1}H NMR spectrum of a bovine plasma sample stored at -20°C for 2 years (top panel) in comparison with a plasma sample stored at the same temperature for 15 years (bottom panel). Selected metabolites are labeled.

Fig. S-2: Scatter plot, trendline and R^2 value for the bucket at 1.34ppm, identified as PC1 when only the residual water signal was excluded from analysis. The bucket contains the CH_3 doublet of lactate. It is obvious that the correlation between bucket integrals and storage time is a negligible one.

Fig. S-3: Scatter plots, trendlines and R^2 values for the buckets contributing to PC1 when the residual water signal and buckets containing lactate signals were excluded from analysis.

Fig. S-4: Scatter plots, trendlines and R^2 values for all buckets with $|r| > 0.5$ (*i.e.* $R^2 > 0.25$; as listed in Tab. 1) when only the residual water signal was excluded from analysis.

Tables

Table S-1: Integral boundaries and relative changes in integral intensities. Calculation of the values as follows: (integral intensity) - (integral intensity of the sample stored for the shortest time) / (integral intensity of the sample stored for the shortest time) x 100 = change in %. Average values and standard errors are listed for three groups (samples stored for less than ten years, stored for ten to eleven years, and stored for more than eleven years). Integral 1 (1.39-4.03ppm) is the integral used for normalization and therefore not listed here.

	boundaries		< 10y		10-11y		> 11y	
	ppm	ppm	avg	SE	avg	SE	avg	SE
Integral 2	8.47	8.44	86.13	67.75	151.15	56.51	183.48	34.25
Integral 3	7.73	7.70	-29.61	14.88	-61.26	2.42	-48.13	8.45
Integral 4	7.70	7.65	14.29	4.93	-7.14	5.37	9.63	17.51
Integral 5	7.39	7.17	1.32	7.98	-8.68	7.04	1.32	23.56
Integral 6	7.03	6.97	-42.40	19.97	-69.97	2.58	-55.52	9.36
Integral 7	6.97	6.92	5.21	4.54	-20.25	5.62	-6.24	16.23
Integral 8	6.87	6.80	-29.33	10.69	-37.24	5.75	-40.53	7.15
Integral 9	5.40	5.37	-28.88	2.92	-46.08	3.96	-44.79	15.77
Integral 10	5.26	5.22	201.84	189.39	644.58	103.80	245.70	197.48
Integral 11	4.18	4.16	-16.24	4.49	9.90	4.47	8.29	7.63
Integral 12	4.15	4.09	-23.35	12.29	-28.36	7.70	-21.41	13.50

Integral 13	4.09	4.04	5.68	12.67	27.07	7.59	50.71	23.27
Integral 14	3.96	3.94	37.32	14.25	12.09	4.52	29.91	13.78
Integral 15	3.94	3.93	4.68	9.29	1.40	2.75	20.78	11.84
Integral 16	3.93	3.91	40.88	21.50	70.98	12.24	42.57	29.52
Integral 17	3.91	3.90	30.76	26.53	−10.84	4.99	−8.34	5.73
Integral 18	3.90	3.89	41.21	21.49	64.72	11.27	40.66	26.29
Integral 19	3.85	3.81	27.80	22.83	73.16	12.20	37.59	28.12
Integral 20	3.80	3.76	33.82	14.72	67.59	9.28	60.87	14.28
Integral 21	3.76	3.70	31.20	29.97	82.14	12.36	25.55	30.04
Integral 22	3.66	3.63	9.66	2.06	4.54	3.00	35.47	15.15
Integral 23	3.59	3.56	13.05	3.24	21.01	4.36	49.73	9.44
Integral 24	3.45	3.38	92.30	86.86	293.16	44.39	139.47	129.81
Integral 25	3.28	3.25	51.85	26.37	−23.52	5.25	−39.30	4.61
Integral 26	3.25	3.21	−29.04	5.95	−25.76	4.68	−43.23	7.41
Integral 27	3.21	3.19	−21.68	14.10	33.18	12.30	59.61	50.58
Integral 28	3.16	3.14	−46.59	14.24	−33.31	22.50	−61.92	12.21
Integral 29	3.05	3.02	−5.70	5.79	−20.17	2.88	−0.66	11.18
Integral 30	3.02	2.92	−3.95	9.64	−24.24	3.31	−9.24	4.99
Integral 31	2.74	2.72	−9.42	2.35	−10.84	2.66	4.47	6.74
Integral 32	2.71	2.63	−11.30	9.82	−14.17	5.15	−0.35	8.82
Integral 33	2.58	2.51	−13.92	12.17	−17.05	5.96	0.08	13.99

Integral 34	2.44	2.41	-28.12	4.53	-15.67	4.06	-6.21	10.81
Integral 35	2.41	2.40	-1.24	13.64	13.70	4.75	33.41	24.03
Integral 36	2.34	2.27	-2.46	9.73	0.29	5.21	22.80	7.09
Integral 37	2.24	2.22	6.72	24.99	-40.30	5.35	-43.79	5.65
Integral 38	2.15	1.96	5.82	4.26	-11.49	3.50	5.67	7.92
Integral 39	1.95	1.89	22.42	16.38	55.06	30.74	18.47	18.21
Integral 40	1.87	1.75	-25.07	11.24	-39.26	4.47	-26.13	10.59
Integral 41	1.75	1.62	-21.89	9.31	-29.95	3.91	-19.12	4.36
Integral 42	1.48	1.41	-13.32	17.99	-27.12	5.61	-21.63	12.10
Integral 43	1.37	1.29	-25.42	14.13	-35.90	6.71	-29.45	14.22
Integral 44	1.24	1.18	-38.20	4.61	-33.81	4.55	-27.06	7.30
Integral 45	1.04	0.99	-37.76	13.88	-55.96	2.11	-48.11	10.25
Integral 46	0.89	0.81	-39.99	7.81	-51.63	3.20	-49.41	8.27

Table S-2: Integrals with correlation to storage time significantly different from random (critical cutoff value: 0.423), and their corresponding significance values and coefficients of determination. The columns 'boundaries' give the upper and lower limits (in ppm) of the integration range for each integral.

integral number	**r**	p	**R^2**	assignment
Int. 25	-0.688	<0.001	0.474	betaine

Int. 37	-0.661	0.001	0.437	acetoacetate
Int. 3	-0.562	>0.05	0.315	imidazole ring
Int. 9	-0.561	0.001	0.315	unidentified singlet
Int. 23	0.552	>0.05	0.305	glycerol
Int. 46	-0.551	0.001	0.303	lipids
Int. 6	-0.490	0.006	0.240	imidazole ring
Int. 20	0.466	>0.05	0.217	glucose

r: correlation; p: significance; R^2: coefficient of determination